Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) in press
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Review

https://doi.org/10.1631/jzus.B2500181

Applications and advances of bile-based omics technologies in the diagnosis of cholangiocarcinoma

Taifeng ZHU[1*], Jiali XING[2*], Ziyue HUANG[1], Shanshan WANG[1], Duo LI[1], Xinting SANG[1], Cong NING[1], Chengpei ZHU[3], Yongchang ZHENG[1✉], Haitao ZHAO[1✉]

[1]*Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College (CAMS & PUMC), Beijing 100730, China*
[2]*State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College (CAMS & PUMC), Beijing 100730, China*
[3]*Department of General Surgery Center, Beijing Youan Hospital, Clinical Center for Liver Cancer, Capital Medical University, Beijing 100069, China*

Abstract: Cholangiocarcinoma (CCA) is challenging to diagnose due to its subtle symptoms and the limited sensitivity and specificity of conventional diagnostic modalities. Therefore, there is an urgent need for more effective detection strategies. Bile, a fluid directly secreted by the biliary system, offers a unique advantage in the diagnosis of CCA by providing a more precise representation of tumor-associated molecules, coupled with its relative ease of collection. The rapid advancement of omics technologies has significantly facilitated the investigation of bile samples. This review explores the applications and advancements of bile-based omics in the diagnosis of CCA from multiple perspectives, including genomics, epigenomics, transcriptomics, proteomics, metabolomics, and microbiomics, aiming to provide new theoretical foundations and technical support for the diagnosis of this malignancy.

Key words: Cholangiocarcinoma; Bile; Omics; Diagnosis

Unedited

1 Introduction

Cholangiocarcinoma (CCA) is a malignant tumor originating from the epithelial cells of the bile ducts, with adenocarcinoma being the predominant histological subtype. Based on the anatomical location of the bile ducts, CCA can be classified into intrahepatic cholangiocarcinoma (iCCA), perihilar cholangiocarcinoma (pCCA), and distal cholangiocarcinoma (dCCA), each subtype differing in clinical characteristics, risk factors, and treatment approaches(Tyson and El-Serag, 2011; Razumilava and Gores, 2013; Ebata et al., 2017). Globally, the incidence of CCA shows significant regional heterogeneity. In Western countries, the incidence remains relatively low (0.1–2/100,000), whereas in Southeast Asia and certain regions of China, it can reach 113/100,000, possibly due to the high prevalence of parasitic infections and chronic cholangitis(Ebata, et al., 2017; Rodrigues et al., 2021). Furthermore, certain chronic biliary diseases, particularly primary sclerosing cholangitis (PSC), have been confirmed as significant risk factors for CCA. Compared to the general population,

✉ Haitao ZHAO, zhaoht@pumch.cn
✉ Yongchang ZHENG, zhengyongchang@pumch.cn
* The two authors contribute equally to this work
Haitao ZHAO, https://orcid.org/0000-0002-3444-8044
Yongchang ZHENG, https://orcid.org/0000-0002-5916-2392

Received Apr. 10, 2025; Revision accepted Oct. 4, 2025;
Crosschecked xxx. xx, 20xx; Published online xxx. xx, 20xx

© Zhejiang University Press 2025

patients with PSC have a 10 to 1000 times higher risk of developing CCA. The pathogenesis is attributed to the persistent inflammation of bile duct cells, leading to abnormal proliferation of these cells, which eventually progresses to cancer(Song et al., 2019). PSC has a significant association with inflammatory bowel disease (IBD), especially ulcerative colitis (UC), with 60% to 80% of PSC patients also having UC or Crohn's disease(Conrad et al., 2014; Song, et al., 2019). Currently, R0 resection, defined as complete surgical removal with negative margins, represents the sole potentially curative intervention for CCA. However, the insidious onset of symptoms often leads to diagnosis at advanced stages, with fewer than 30% of patients indicated for surgery (Joo et al., 2018; Gunasekaran et al., 2020). For patients with advanced-stage CCA deemed ineligible for surgery, the gemcitabine and cisplatin combination chemotherapy regimen constitutes the current standard care. While this regimen may confer a modest delay in tumor progression and palliation of symptoms, the inherent chemoresistance of CCA limits overall efficacy, resulting in a five-year survival rate of 2-5%(Razumilava and Gores, 2013; Harding et al., 2023). In recent years, with the rise of molecular targeted therapies and immunotherapies, precision medicine has brought new hope to the treatment of CCA. However, currently only a subset of patients can benefit from these treatments and most targeted therapies are still in clinical trials, with their efficacy requiring further validation (Valle et al., 2021; Harding, et al., 2023).

Early diagnosis is a critical determinant of survival in patients with CCA. However, the insidious nature of clinical symptomatology and the complex anatomical localization of the tumor render traditional serum markers, such as carbohydrate antigen 19-9 (CA 19-9) and conventional imaging modalities, suboptimal with respect to diagnostic sensitivity and specificity(Rösch et al., 2002; Nehls et al., 2004; Sinakos et al., 2011; Joo, et al., 2018). In addition, other biliary diseases may potentially interfere with the expression of biomarkers. For instance, choledocholithiasis can cause biliary obstruction and an inflammatory response, which may lead to an increase in the expression of biomarkers associated with biliary damage, such as bile duct-specific proteins and inflammatory factors, thereby affecting their specificity(Bombaci et al., 2019). PSC and UC may also result in elevated levels of certain biomarkers, such as CA 19-9, through chronic inflammation and fibrosis processes(Wannhoff et al., 2015). Endoscopic ultrasound-guided fine-needle aspiration, in conjunction with endoscopic retrograde cholangiopancreatography (ERCP) and brush cytology, has gradually been applied to the diagnosis of biliary malignancies. However, their diagnostic sensitivity is relatively low, with potential severe complications reported(Glasbrenner et al., 1999; Byrne et al., 2004; De Vries et al., 2019). These limitations have prompted researchers to explore safer and more efficient diagnostic strategies, including bile-based testing. Owing to its direct contact with the neoplastic lesion in CCA, bile constitutes an ideal biofluid for tumor biomarker analysis. Studies have shown that, compared to blood and other bodily fluids, bile contains higher concentrations of tumor-derived molecules, such as cell-free DNA (cfDNA), microRNAs (miRNAs), proteins, and metabolites, thereby significantly enhancing detection sensitivity(Gaetano et al., 2011; Cote et al., 2014; Li et al., 2022c). Furthermore, bile has a distinct metabolic composition, with a narrower metabolic range than the systemic circulation. This reduces the dilution or interference of systemic metabolic processes on tumor molecule concentrations, making CCA-related molecules more easily detectable in bile(Zhang et al., 2023a; Yang et al., 2025). Moreover, changes in the diversity and composition of the biliary microbiota have been implicated in the development of CCA, providing a new research direction for its diagnosis(Han et al., 2020; Li, et al., 2022c). Most importantly, the collection of bile samples is clinically feasible. Specifically, during procedures such as endoscopic retrograde cholangiopancreatography (ERCP) or percutaneous transhepatic biliary drainage (PTBD), bile samples can be obtained directly during the operation without the need for additional invasive procedures, significantly reducing the burden on patients(Ke et al., 2018; Zhang, et al., 2023a). This minimally invasive and efficient approach makes bile samples highly promising for clinical diagnostic applications.

In recent years, the rapid development of omics technologies has led to unprecedented breakthroughs in tumor research and clinical applications. These technologies have provided in-depth insights into the mechanisms of tumor initiation and progression at the molecular level, demonstrating tremendous potential in

Uncorrected

diagnosis, personalized treatment, and prognosis assessment of tumors(Chang et al., 2020; Akhoundova and Rubin, 2022; Tang et al., 2024). This review comprehensively explores the diagnostic value of bile samples in CCA across multiple omics domains, including genomics, epigenomics, transcriptomics, proteomics, metabolomics, and microbiomics (Fig. 1). A comprehensive literature search was conducted using several databases, including PubMed and Web of Science. The following search terms were used: cholangiocarcinoma/bile duct cancer, bile, biomarker/biosignature/molecular marker and diagnosis. Studies that did not use omics technologies were excluded. The full text of potentially relevant studies was reviewed.

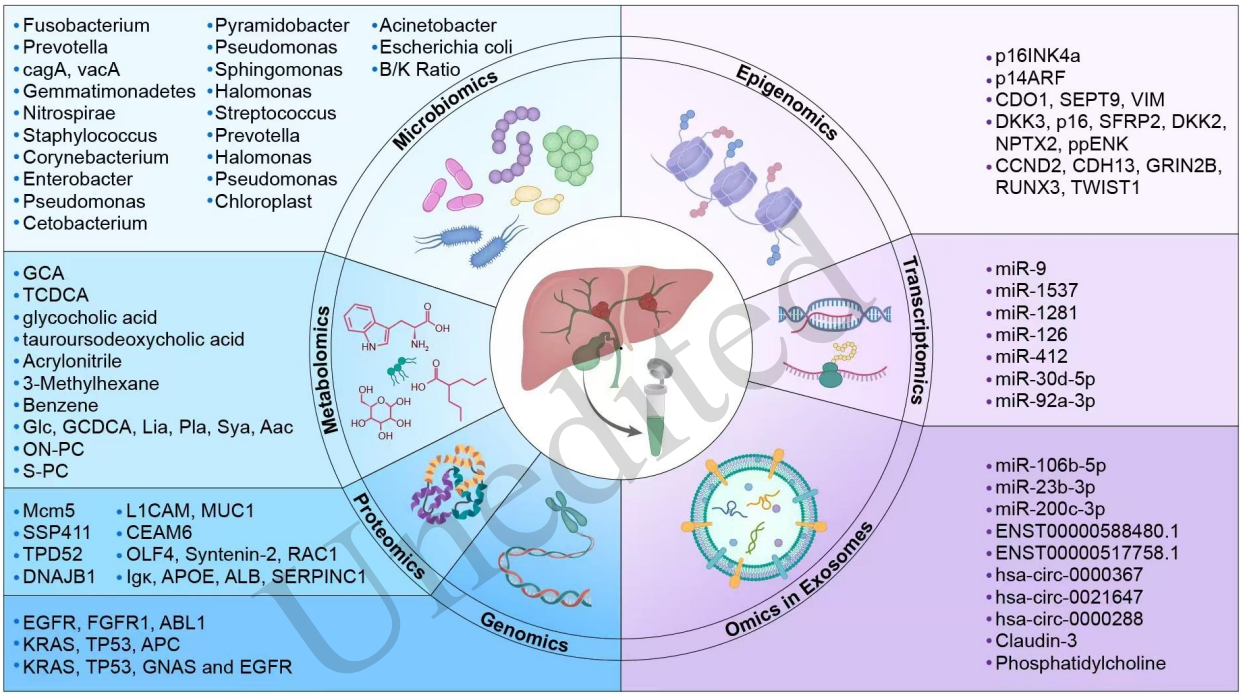

Fig. 1 Applications of Bile-based Omics Technologies in the Diagnosis of Cholangiocarcinoma. Created with BioRender.com.

2 Bile-based genomics and epigenomics

Whole-genome sequencing (WGS) technology has become a pivotal tool in genomics, particularly for the identification of key gene mutations that drive tumorigenesis. This methodology enables comprehensive analysis of the entire genome, facilitating the classification of tumors and the discovery of potential therapeutic targets(Kim, 2013; Emmanuel et al., 2014; Panea et al., 2019). Epigenomics plays a crucial role in cancer research by identifying tumor-specific epigenetic markers through the analysis of various epigenetic modifications. These modifications, including DNA methylation, histone modifications, and alterations in chromatin structure, can contribute to the differentiation of tumor subtypes and facilitate early diagnosis(Maury and Hashizume, 2017; Luo et al., 2020; Cao et al., 2022).

2.1 cfDNA

cfDNA comprises DNA fragments released from cells into blood or other bodily fluids, which can serve as a biomarker for diseases, particularly cancer. This utility stems from the capacity of cfDNA to reflect the genomic variation and gene expression profiles of tumor cells, thereby providing crucial information for tumor diagnosis and monitoring (Jahr et al., 2001; Siravegna et al., 2017; Ptashkin et al., 2018). Li et al. (2022) found that the concentration of cfDNA in bile was significantly higher than in plasma (median values of 1918 ng/ml

and 63.1 ng/ml, respectively, $p = 0.0017$, Fig. 2a). High-throughput sequencing revealed missense mutations as the predominant type of mutation, with a significantly higher number of missense mutations observed in bile cfDNA than in plasma (Figs 2b and 2c). Furthermore, genomic mutations identified in bile cfDNA showed a high degree of concordance with those in tumor tissue (Fig. 2d). These results suggest that bile cfDNA may serve as a potential surrogate for tumor tissue biopsy(Li, et al., 2022c). Lee et al. (2016) conducted a comparative genomic analysis of gene mutations in 24 bile samples and 17 tumor tissue samples. The study identified significant mutations in genes such as *EGFR* and *PIK3CA* in bile from patients with infiltrative CCA, while *TP53* and *KRAS* represented the primary mutated genes in mass-forming CCA. This suggests that the two subtypes may represent distinct tumor entities, which could help further understand their pathological features and suggest potential therapeutic targets (Lee et al., 2016). Arechederra et al. (2022) developed the next-generation sequencing (NGS)-based Bilemut assay (where "Bilemut" refers to a mutational analysis method based on bile cfDNA) and performed bile cfDNA mutation analysis on 68 patients with biliary strictures. The results showed a sensitivity of 96.4% in detecting malignant strictures, with a sensitivity of 100% in cases initially classified as benign or indeterminate strictures. These findings suggest that Bilemut detection can provide significant clinical value in the early diagnosis of biliary diseases, particularly in the accurate identification of malignant lesions(Arechederra et al., 2022). Shen et al. (2022) developed an efficient cfDNA extraction method based on a silicon membrane (3D-BCF, Fig. 3a) to address the issue of low recovery efficiency of long fragments of cfDNA in bile using commercial kits. The experimental results showed that this method not only significantly improved cfDNA extraction purity (i.e., the quality and integrity of cfDNA extracted from bile samples) but also outperformed other kits in coverage across short, medium, and long DNA fragments, providing technical support for the precise detection of bile cfDNA(Shen et al., 2022). Collectively, the diagnostic potential of bile cfDNA in CCA has been validated by numerous studies. These studies not only demonstrate the high diagnostic sensitivity and specificity of bile cfDNA but also highlight significant improvements in analytical efficiency through technological advancements, thereby laying the foundation for the advancement of precise diagnosis and personalized treatment of biliary cancers.

Unedited

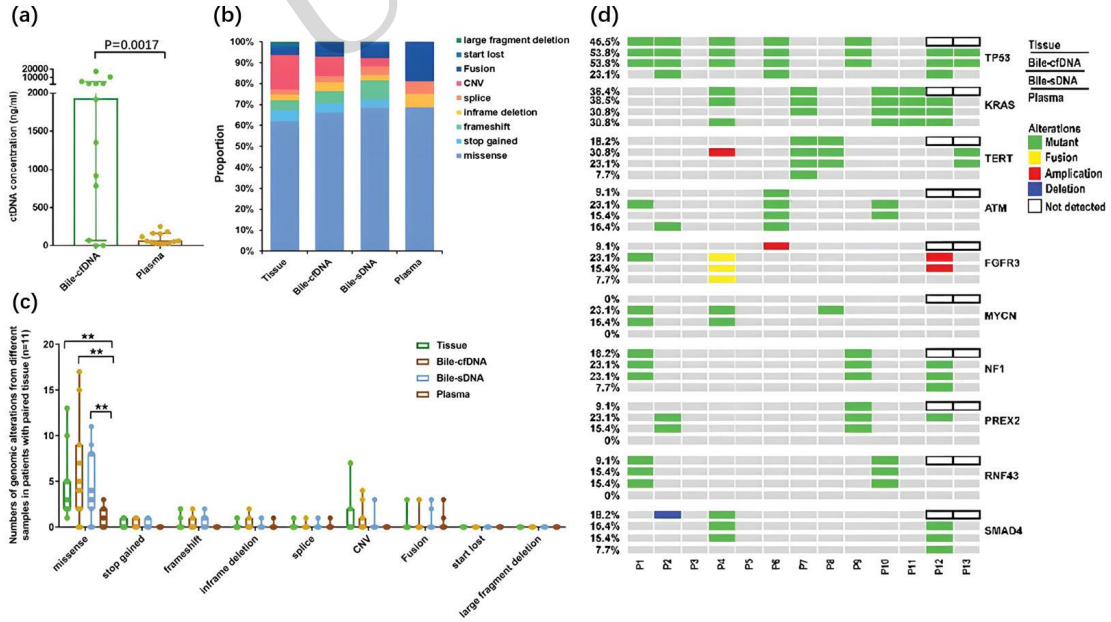

Fig. 2 Bile is a reliable and valuable source for studying cfDNA in cholangiocarcinoma. (a) Comparison of cfDNA concentrations in 11 paired bile supernatants and plasma samples. (b) The proportion of genomic alteration types in different samples from patients with paired tissues. (c) The number of various genomic alterations in different samples from patients with paired tissues. **p-value < 0.01. (d) Profiles of recurrent gene genomic alterations in different sample types. (A–D) Reprinted from Li et al., 2022.

2.2 DNA methylation

In recent years, DNA methylation markers in bile, a focal point of epigenomic research, have contributed to significant progress in the application of diagnosis to CCA(Nakaoka et al., 2017; Wasenang et al., 2019). The fundamental principle of this method lies in the detection of aberrant methylation patterns within CpG islands in gene promoter regions, as these abnormalities often lead to the silencing of tumor suppressor genes, which is a key pathogenic mechanism in malignant tumors(Herman et al., 1995; Chin et al., 1998). For example, the *p16INK4a* and *p14ARF* genes encoded by the INK4a/ARF locus regulate the cell cycle by inhibiting RB protein phosphorylation and stabilizing p53, respectively. The methylation status of their promoters has become an important target for molecular diagnosis(Merlo et al., 1995; Quelle et al., 1995; Chin, et al., 1998). A pioneering study proposed the use of bile samples to detect methylation of the INK4a/ARF promoter as a non-invasive molecular diagnostic method. The results revealed methylation detection rates of 52.2% for p16INK4a and 47.6% for p14ARF in patients with CCA, significantly higher than those in patients with non-malignant diseases(Klump et al., 2003). Zhang et al. (2010) analyzed the methylation status of 19 tumor suppressor genes in 80 bile samples and identified a six-gene panel, including *DKK3*, *p16*, *SFRP2*, *DKK2*, *NPTX2*, and *ppENK*. When a methylation index (MI) threshold of 0.5 was applied, this panel showed a sensitivity of 77.27% and a specificity of 77.78% in diagnosing malignant biliary strictures(Zhang et al., 2010). In contrast, Shin et al. (2012) optimized a five-gene panel (*CCND2*, *CDH13*, *GRIN2B*, *RUNX3*, and *TWIST1*) through the screening of 59 DNA methylation markers, demonstrating significant diagnostic efficacy for extrahepatic cholangiocarcinoma (EHC) in a validation cohort. The panel achieved a sensitivity of 83% and a specificity of 100%, significantly surpassing the sensitivity of conventional bile cytology (46%). Moreover, the study emphasized the stability of methylation detection, which showed high reproducibility and accuracy, even when performed on bile samples(Shin et al., 2012). Vedeld et al. (2021) further used digital PCR technology to detect the methylation status of genes such as *CDO1*, *CNRIP1*, *SEPT9*, and *VIM*, and found that the combined detection of these genes yielded an area under the curve (AUC) of 0.88 in the diagnosis of primary sclerosing cholangitis (PSC)-associated CCA, with a sensitivity of 79% and a specificity of 90%. Notably, this method was able to detect CCA up to 12 months prior to clinical diagnosis, providing important technical support for early diagnosis(Vedeld et al., 2021). From a technical perspective, bile methylation analysis combined with multiple genetic markers significantly enhances diagnostic efficacy, overcoming the limitations of traditional cytological methods caused by insufficient shed cell counts(Ferrari Júnior et al., 1994; Kipp et al., 2004). Concomitantly, high-sensitivity detection technologies, such as digital PCR and MethyLight, have enhanced the accuracy and reproducibility of methylation detection(Shin, et al., 2012; Abe et al., 2019). However, current studies are focused mostly on specific populations with limited sample sizes, and the generalizability and broader application of their results still need further validation.

3 Bile-based transcriptomics

Transcriptomics, a high-throughput analytical technology centered on RNA, has become a transformative tool in understanding gene expression dynamics and its implications for diseases, particularly cancer. By leveraging advancements in RNA sequencing (RNA-seq) and related methodologies, transcriptomics has shown significant promise in the accurate diagnosis and treatment of cancers(Sager et al., 2015; Ergin et al., 2022; Satam et al., 2023).

MicroRNAs constitute a class of small, non-coding RNAs that play a key role in the post-transcriptional regulation of gene expression and are involved in various biological processes, including the development and progression of cancer(Cummins and Velculescu, 2006; Meng et al., 2006; Chen et al., 2009). In CCA, a highly invasive biliary tumor, miRNAs have attracted significant interest due to their potential use in early diagnosis. Studies have shown that specific miRNAs (e.g., miR-21, miR-9, and miR-26a) show good sensitivity and

Uncorrected

specificity across various sample types, including tissues, bile, and serum. Notably, bile-derived miRNAs demonstrate superior diagnostic performance compared to miRNAs derived from other sample sources(Liang et al., 2016). Gaetano et al. (2011) screened miR-9 from 667 miRNAs and validated its exceptional diagnostic efficacy in bile, yielding an AUC of 0.88, a sensitivity of 88.9%, and a specificity of 100%. Furthermore, the long-term stability of miR-9 in bile (i.e., maintenance of activity under storage and transportation conditions) further highlights its potential for clinical diagnostic applications(Gaetano, et al., 2011). However, a study by Syn et al. (2015) did not identify a significant correlation of miR-9 in patients with PSC and/or CCA. This discrepancy may be attributed to the lack of technical standardization (e.g., differences in RNA extraction and analysis methods), variation in sample quality (such as bile collection time and storage conditions), and patient population characteristics (e.g., bile composition in PSC patients may be influenced by chronic inflammation). The results of this study indicated that a model incorporating miR-1537 and CA19-9 achieved an AUC of 0.91, with a specificity of 93%, showing that the combined analysis of miRNAs and traditional markers enhances diagnostic performance (Syn et al., 2015). Han et al. (2020) screened miR-30d-5p and miR-92a-3p from bile and observed that their expression levels were significantly elevated in CCA patients compared to those with benign biliary diseases (BBD). Quantitative real-time polymerase chain reaction (qRT-PCR) validation showed that miR-30d-5p exhibited favorable diagnostic capability, yielding an AUC of 0.730, a sensitivity of 81.1%, and a specificity of 60.5%, outperforming traditional serum markers CA19-9 (AUC = 0.675) and CEA (AUC = 0.603). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that miR-30d-5p is implicated in CCA signaling pathways closely associated with cell proliferation, differentiation, and apoptosis, such as the Hippo, Wnt, and p53 pathways, while miR-92a-3p is enriched in the mitogen-activated protein kinase (MAPK) and epidermal growth factor receptor (EGFR) pathways, which have been shown to play critical roles in the initiation and progression of CCA. In particular, KEGG pathways involved in key processes such as ECM degradation, cell adhesion, and angiogenesis play a crucial role in the ability of tumor cells to breach local tissue barriers, enter the bloodstream, and form distant metastatic sites. These findings provide mechanistic support for their potential as diagnostic biomarkers(Han, et al., 2020). Despite their promising potential, miRNAs are not yet standardized for clinical use, but their diagnostic value could be further enhanced when combined with other omics approaches, providing a more robust and comprehensive tool for CCA diagnosis.

Unedited

4 Bile-based proteomics

Proteomics, a discipline dedicated to the comprehensive study of protein composition, structure, function, and interactions, represents a cornerstone of post-genomic research(Zhang et al., 2013; Cui et al., 2022). The advent and rapid advancement of high-throughput technologies have facilitated a paradigm shift in proteomics research, transitioning from conventional methodologies to more precise and efficient approaches, driven by the implementation of emerging technologies, including mass spectrometry and liquid biopsy(Gao et al., 2023; Lapitz et al., 2023). The first large-scale study of bile proteomics was conducted in 2004, when 87 proteins were identified in the bile of healthy patients(Kristiansen et al., 2004). Subsequently, contemporary investigations leveraging advanced technologies have delineated a more extensive repertoire of bile proteins (Lukic et al., 2014; Ren et al., 2019). Within the domain of tumor diagnostics, proteomics provides crucial support for early diagnosis and precision treatment of tumors by identifying and analyzing protein markers in bodily fluids such as serum and bile.

4.1 Discovery of novel biomarkers

Pioneering research by Ayaru et al. (2008) initially explored the diagnostic value of minichromosome maintenance protein 5 (MCM5) in bile, revealing significantly elevated levels in the bile of CCA patients

compared to benign cases. The detection sensitivity of MCM5 was 66%, with a specificity of 97%, significantly outperforming conventional brushing cytology methods (sensitivity of 20%)(Ayaru et al., 2008). This finding provided a direction for the screening of bile protein biomarkers, laying the foundation for conducting CCA diagnostic research using proteomics technologies. Subsequently, Mulvenna et al. (2012) identified 49 potential bile biomarkers using two-dimensional electrophoresis-mass spectrometry (2-DE/MS), among which SSP411 demonstrated superior sensitivity and specificity compared to the traditional marker CA19-9, showing excellent diagnostic performance(Mulvenna et al., 2012). With the advancement of high-throughput proteomics technologies, Ren et al. (2019) used isobaric tags for relative and absolute quantitation (iTRAQ) coupled with liquid chromatography-tandem mass spectrometry (LC-MS/MS) to identify 778 differentially expressed proteins in bile samples, ultimately validating TPD52 and DNAJB1 as key biomarkers for CCA. These biomarkers not only showed significantly elevated levels in bile but were also associated with the clinical prognosis of patients(Ren, et al., 2019). The implementation of glycoproteomics has further refined the precision of biomarker research. Matsuda et al. (2013) used lectin wheat germ agglutinin (WFA) enrichment to screen glycosylated proteins in bile and identified L1 cell adhesion molecule (L1CAM) as a CCA biomarker with a specificity of 93%. The combined analysis of L1CAM and WFA-positive sialyl MUC1 significantly improved diagnostic accuracy (AUC = 0.93). Although glycoprotein biomarkers demonstrate high specificity, their technical complexity limits their widespread application(Matsuda et al., 2013). Besides, Son et al. (2020) conducted a comprehensive analysis of the bile proteome in patients with EHC using LC-MS/MS, identifying 245 differentially expressed proteins. The significant upregulation of immunoglobulin κ light chain and apolipoprotein E suggests novel avenues for potential biomarkers (Son et al., 2020). However, current studies still face issues such as limited sample sizes, regional limitations, and a lack of multi-center validation.

4.2 Validation of known biomarkers

CEAM6, an important member of the immunoglobulin superfamily, has been shown to be significantly overexpressed in various tumor tissues, including breast cancer, pancreatic cancer, and iCCA, and is strongly associated with the invasiveness and metastatic potential of neoplastic cells(Duxbury et al., 2005; Ieta et al., 2006; Blumenthal et al., 2007). Specifically, in pancreatic cancer, its expression level correlates not only with the degree of differentiation, lymph node metastasis, and pathological staging, but also serves as a potential prognostic indicator(Duxbury et al., 2004; Duxbury, et al., 2005). However, despite the extensive research on CEAM6 in tissue samples, its application in liquid biopsy, particularly in bile samples, and its diagnostic potential remain underexplored(Farina et al., 2009; Farina et al., 2011). Farina et al. (2014) further confirmed that the elevated expression of CEAM6 in bile samples could differentiate malignant from benign biliary strictures, with an AUC of 0.92, sensitivity of 93%, and specificity of 83%. Moreover, the diagnostic efficiency was further improved (AUC = 0.96) when CEAM6 was combined with the serum marker CA19-9, demonstrating the potential of composite biomarker application(Farina et al., 2014). Currently, CEAM6 remains in the investigational phase and has not been validated for clinical use. Ongoing research is focused on its further validation, and larger-scale, multi-center studies are needed to assess its clinical applicability.

4.3 Development and optimization of diagnostic techniques

Bile proteomics research has progressively shown enhanced diagnostic value with the advancement and optimization of analytical technologies. Historically, the complex chemical composition of bile, characterized by elevated levels of lipids and salts, has posed substantial challenges to proteomics research by impeding protein extraction and identification. Lukic et al. (2014) mitigated interference from these substances in bile through differential centrifugation combined with iTRAQ labeling and LC-MS/MS, resulting in a five-fold increase in the number of proteins identified compared to conventional methodologies (Fig. 3b). In malignant bile samples, 104 overexpressed cancer-related proteins were identified for the first time, including olfactomedin-4 (OLF4), syntenin-2, and Rac1. These proteins are believed to be closely associated with tumor

Unedited

cell invasion and migration, laying a foundation for the development of CCA diagnostic biomarkers(Lukic, et al., 2014). Early CE-MS studies established the feasibility of bile peptidomics to distinguish CCA from PSC and choledocholithiasis with independent-cohort validation(Lankisch et al., 2011). These data provided the first disease-specific bile peptide patterns in a PSC-enriched differential diagnosis scenario and set the stage for subsequent urine- and combined fluid-based classifiers. Complementing bile analysis, a CE-MS based urine proteomic classifier differentiated CCA from PSC and other benign biliary disorders in cross-sectional validation (AUC 0.87; 83% sensitivity; 79% specificity), thereby enabling a non-invasive approach that informed later combined bile–urine models(Metzger et al., 2013). Voigtländer et al. (2017) integrated bile and urine proteomics data to develop a composite diagnostic model based on logistic regression. In a retrospective study involving 87 patients, the model achieved a sensitivity of 92% and a specificity of 84%, with an AUC of 0.96. In a prospective validation cohort of 45 patients, the sensitivity further increased to 94%, significantly outperforming the analysis of bile (63%) or urine alone (81%). This model provides critical guidance for the precise screening of high-risk populations for CCA, such as patients with primary sclerosing cholangitis, particularly in instances where conventional imaging modalities (e.g., MRI and CT) encounter difficulties in differentiating between malignant and benign biliary strictures (Voigtländer et al., 2017). Building upon these findings, the research team further explored the mechanisms of bile and urine proteomic biomarkers using protease profiling and gene ontology (GO) analysis. The results showed that the activity of ADAMTS4 protease in bile was significantly elevated. As a matrix metalloproteinase, ADAMTS4 plays a key role in the epithelial-mesenchymal transition (EMT). This process is crucial for tumor cell migration, invasion, and metastasis. ADAMTS4 contributes to these processes by degrading the extracellular matrix (ECM) and facilitating the EMT, thereby supporting tumor progression. These mechanisms not only reveal local pathological changes reflected in bile biomarkers but also mirror systemic biological alterations through urine biomarkers, providing novel perspectives and important insights for the molecular mechanism research of CCA(Voigtländer et al., 2020).

Uncorrected

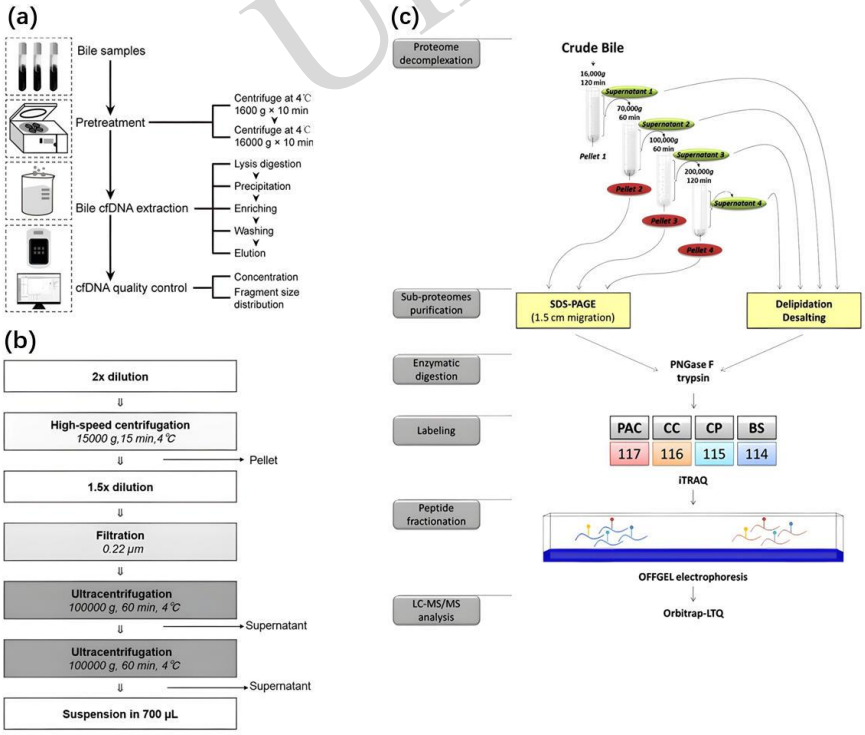

Fig. 3 Improvements in bile sample processing techniques. (a) Bile cfDNA extraction procedure and cfDNA quality control of the 3D-BCF method. Reprinted from Shen et al., 2022. (b) Bile fractionation workflow based on differential

centrifugation. Reprinted from Lukic et al., 2013. (c) Workflow for efficient isolation of high-purity exosomes from human bile. Reprinted from Ikeda et al., 2021.

5 Bile-based metabolomics

Metabolomics, a systems biology approach, elucidates alterations in metabolic pathways and their roles in disease pathogenesis by analyzing small molecular metabolites within biological organisms(Tomacha et al., 2021; Wang et al., 2023). In recent years, metabolomics technologies, particularly mass spectrometry-based analytical methodologies such as LC-MS and gas chromatography-mass spectrometry (GC-MS), have emerged as pivotal tools in oncological research, owing to their high sensitivity and selectivity(Cai et al., 2023; Wang, et al., 2023; Jackson et al., 2024). Using these technologies, researchers have not only discovered highly specific biomarkers for the diagnosis of CCA but have also advanced the development of precision treatment strategies, laying the foundation for their clinical application.

5.1 Bile acid metabolites

Bile acids are important metabolites in bile, produced mainly by the liver through cholesterol metabolism. They can be divided into primary bile acids (such as chenodeoxycholic acid, CDCA), which are directly generated by the liver, and secondary bile acids (such as lithocholic acid, LCA), which are generated through the conversion of primary bile acids by gut microbiota(Russell, 2003; Wahlström et al., 2016). Beyond their pivotal role in lipid digestion, bile acids function as signaling molecules, regulating the metabolic homeostasis of the liver and intestines(Thomas et al., 2008). In the microenvironment of CCA, bile acid metabolism may be dysregulated, making it an important direction for investigating its diagnostic value. Song et al. (2018) demonstrated a significant increase in the proportion of glycocholic acid (GCA) in the bile of CCA patients, reaching 35.6%, markedly higher than that observed in benign biliary diseases (22.3%, $p<0.0001$) and pancreatic adenocarcinoma (19.9%, $p<0.0001$). In contrast, the proportion of taurine-conjugated chenodeoxycholic acid (TCDCA) was significantly reduced, registering only 7.3%. The increase in GCA and reduction in TCDCA levels have been linked to alterations in key signaling pathways, such as the Wnt/β-catenin and MAPK pathways, which are involved in cell proliferation and tumor progression. Functional validation showed that GCA significantly upregulated the expression of TGR5 and S1PR2 genes in CCA cells, thereby promoting tumor cell proliferation. In contrast, TCDCA did not have similar effects, possibly due to its weaker ability to activate the related signaling pathways (Song et al., 2018). Furthermore, Zhang et al. (2023) observed analogous metabolic abnormalities in patients with pCCA. Metabolites such as GCA and TCDCA showed significant changes in bile acid metabolism and bile secretion-related pathways and were correlated with serum total bilirubin (TBIL) and CA19-9 levels, further substantiating their potential value as diagnostic biomarkers for CCA (Zhang, et al., 2023a).

5.2 Lipid metabolites

Lipid metabolites constitute a significant component of bile, originating from cholesterol metabolism, cell membrane turnover, and physiological processes within the liver and biliary system. They play crucial roles in regulating cell signaling, inflammatory responses, and energy metabolism(Giacomini et al., 2021; Vassiliou and Farias-Pereira, 2023). In CCA, due to changes in the tumor microenvironment, the composition and levels of specific lipid metabolites are altered, manifested by their abnormal accumulation or depletion (Giacomini, et al., 2021; Padthaisong et al., 2021). These metabolic shifts can be detected using techniques such as metabolomics, further validating their potential as biomarkers for the diagnosis of CCA(Li et al., 2022a). Navaneethan et al. (2014) analyzed bile samples from 46 patients with biliary diseases using liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) and found that the levels of oxidized nonyl-acyl phosphatidylcholine (ON-PC) and succinyl-phosphatidylcholine (S-PC) were significantly higher in CCA patients than in those with benign biliary diseases. The combined detection of ON-PC and S-PC yielded an AUC of 0.91, with a sensitivity of 100% and a specificity of 83.3%, demonstrating exceptional diagnostic

Unedited

performance(Navaneethan et al., 2014a). In another study, Yang et al. (2025) used nanoparticle-enhanced laser desorption/ionization mass spectrometry (NPELDI-MS) technology to construct a bile metabolic fingerprint (BMF). They found that the expression of several bile acids and lipid metabolites was significantly reduced in tumor patients, while a subset, including glycochenodeoxycholic acid (GCDCA), showed elevated levels in cancer patients. These metabolic changes suggest that bile composition may be closely related to the occurrence and development of CCA(Yang, et al., 2025). In contrast, the study by Navaneethan et al. (2014) provided more detailed data, validating the high efficiency of combined detection of ON-PC and S-PC in diagnosis, while the study by Yang et al. (2025) emphasized the technological advantages. Lipids serve as complementary biomarkers to bile acids in cancer diagnosis, with the potential for synergistic roles in diagnostic panels.

5.3 Volatile organic compounds

Volatile organic compounds (VOCs) constitute a class of metabolites characterized by high vapor pressure, ubiquitously present in various bodily fluids such as bile, blood, and exhaled breath. They originate mainly from redox reactions in metabolic processes, lipid peroxidation, and metabolic activities of the gut microbiota(De Meij et al., 2013; Hanouneh et al., 2014; Navaneethan et al., 2014b). Under normal physiological conditions, VOCs participate in cell signaling and metabolic regulation, maintaining metabolic homeostasis within the organism. However, in disease states, the generation and distribution of these metabolites are significantly affected, and their metabolic profiles show specific patterns of change. Therefore, VOCs are considered potential biomarkers for diagnostic and monitoring purposes(Altomare et al., 2013; Navaneethan, et al., 2014b). Navaneethan et al. (2015) analyzed bile samples from 32 patients using selected ion flow tube mass spectrometry (SIFT-MS) and found that the concentrations of acrylonitrile, 3-methylhexane, and benzene were significantly lower in CCA patients compared to patients with PSC alone ($p < 0.05$). A logistic regression model based on these metabolites yielded an AUC of 0.89, with a diagnostic sensitivity of 90.5% and a specificity of 72.7%, demonstrating substantial discriminatory ability. However, the limited sample size of the study (only 11 PSC patients with CCA) may limit the generalizability of its conclusions(Navaneethan et al., 2015). In contrast, Gui et al. (2023) conducted a study with a larger cohort (200 cases) and used advanced gas chromatography-ion mobility spectrometry (GC-IMS) detection technology, identifying 12 significantly differentiated VOCs. Among these, 4 (e.g., 2-ethyl-1-hexanol and cyclohexanone) were upregulated in CCA patients, while 8 aldehydes (e.g., hexanal and (E)-2-hexenal) were downregulated. The study used a support vector machine (SVM) model for diagnostic performance evaluation, achieving an AUC of 0.966, with a sensitivity of 93.1% and a specificity of 100%. These results not only validate the diagnostic potential of VOCs but also highlight the promising application of machine learning methods in metabolomics(Gui et al., 2023). Despite the promising potential of VOCs as biomarkers, standardization and reproducibility remain significant barriers to their clinical application. Overcoming these challenges is crucial for their adoption in clinical practice.

6 Bile-based microbiomics

Microbiomics is a field in which the interactions between the host and the microbiome ecosystem within the body are studied. Using high-throughput sequencing technologies, such as 16S rRNA and metagenomic sequencing, it analyzes the composition, function, and metabolic products of microorganisms. These research findings provide important tools for exploring disease mechanisms and developing diagnostic methods(Lederer et al., 2023; Han et al., 2024). Historically, bile was considered an aseptic fluid, involved mainly in lipid metabolism and waste excretion. However, with the rapid development of molecular biology and high-throughput sequencing technologies, researchers have discovered that bile contains a rich and diverse bacterial community(Binda et al., 2022; Nehring et al., 2023). An increasing number of studies suggest that the development of CCA involves chronic inflammation, metabolic disorders, and immune dysregulation, and these

Unedited

pathological changes may be closely related to the dynamic changes in the bile microbiome. This provides new theoretical support for the early diagnosis of CCA (Mima et al., 2017; Miyabe et al., 2022; Ye et al., 2023).

Helicobacter pylori is a Gram-negative, spiral-shaped bacterium that typically colonizes the gastric mucosa and is widely recognized as being closely associated with the development of chronic gastritis and gastric carcinoma(Apostolov et al., 2009). Studies have detected *H. pylori* DNA and antigenic components in bile, suggesting its potential colonization of the biliary system and contribution to CCA pathogenesis through mechanisms such as the induction of chronic inflammation or the alteration of the biliary microenvironment(Fallone et al., 2003; Apostolov, et al., 2009; Segura‐López et al., 2015). Avilés-Jiménez et al. (2016) used 16S rRNA sequencing and PCR techniques to show that *H. pylori* and its virulence genes (*cagA* and *vacA*) were significantly enriched in the bile of patients with EHC. These strains may promote CCA formation by stimulating the host immune response, inducing bile duct epithelial cell damage, and causing chronic inflammation(Avilés-Jiménez et al., 2016). Other studies have revealed differences in specific microbiota between CCA and benign biliary diseases. Chen et al. (2019) analyzed the bile microbiome of patients with dCCA and choledocholithiasis and found that species from the *Proteobacteria* phylum (such as *Escherichia coli*) were significantly enriched in CCA patients, while *Firmicutes* predominated in choledocholithiasis patients. This significant divergence in microbiota composition suggests that bile microecological imbalance may reflect distinct pathological states and play a pivotal role in CCA development(Chen et al., 2019). In the exploration of potential diagnostic biomarkers, Dangtakot et al. (2021) used linear discriminant analysis (LDA) to identify microbiota differences between CCA and choledocholithiasis patients. Among these, *Enterobacter* and *Pseudomonas* were significantly more abundant in CCA patients, while *Streptococcus* predominated in choledocholithiasis patients. This finding provides a basis for using the bile microbiome in disease differentiation(Dangtakot et al., 2021). Li et al. (2022) analyzed the bile microbiota of patients with pCCA and dCCA and identified the three most abundant bacterial genera in pCCA as *Pseudomonas*, *Sphingomonas*, and *Halomonas*, while in dCCA, they were *Streptococcus*, *Prevotella*, and *Halomonas*. This suggests that distinct CCA subtypes may exhibit unique microbiological ecological features, providing new insights into the pathological classification of the disease(Li et al., 2022b). At the same time, innovative microbiome indicators have provided new directions for early diagnosis. Zhang et al. (2023) proposed a diagnostic method based on the *Bifidobacterium*/*Klebsiella* (B/K) ratio and found that this ratio was significantly decreased in CCA patients, offering a novel tool for non-invasive early detection(Zhang et al., 2023b). Furthermore, Lee et al. (2024) integrated microbiome and metabolome data and discovered that isoleucine levels were significantly elevated in the bile of CCA patients, while *E. coli* was notably enriched. Although the causal relationship between these observations remains unclear, these changes may reflect metabolic disturbances and microbial dysbiosis in CCA. *In vitro* experiments showed that high concentrations of isoleucine significantly inhibited the proliferation of CCA cells without inducing apoptosis. This suggests that isoleucine may play a role by interfering with tumor metabolism, while microbiome changes may influence cancer development by regulating the metabolic environment(Lee et al., 2024).

Research on the bile microbiome is still in its early stages, with challenges such as limited sample sizes, substantial heterogeneity in patient characteristics, and inconsistent study designs, which limit the generalizability of the findings. While current studies have observed shifts in the microbiota in CCA patients, these associations remain correlative and the causality between microbiota changes and disease progression has yet to be established. Future research in microbiome-CCA associations will need to include larger patient cohorts, consistent study designs, and mechanistic studies to transition from correlation to causality in understanding the role of the microbiome in CCA.

7 Omics in bile exosomes

Unedited

Exosomes are phospholipid bilayer vesicles secreted by cells, ranging in size from 30 to 150 nm. They are released into the extracellular space through the fusion of multivesicular bodies with the plasma membrane and are ubiquitously present in various bodily fluids such as blood, bile, and urine. Exosomes carry bioactive molecules, including proteins, nucleic acids (such as miRNA, mRNA, and circular RNA), and lipids, and are capable of transmitting intercellular information, reflecting the physiological and pathological status of the originating cells(Liu et al., 2024; Miceli et al., 2024; Shu et al., 2024). In recent years, with the continuous advancement of efficient separation and analytical technologies, the application of exosomes in early cancer diagnosis has gradually become a research hotspot(Liu, et al., 2024; Wang et al., 2024). Unlike other bodily fluids such as blood, bile is derived directly from the biliary system, and the contents carried by its exosomes better reflect local pathological changes, with higher concentrations of biomarkers, thereby improving the sensitivity and specificity of CCA diagnosis.

7.1 Transcriptomics

Exosomes, functioning as stable biological carriers, provide protection for encapsulated nucleic acid molecules (including miRNA, lncRNA, and circRNA) against degradation by the external environment, rendering them a significant avenue for the investigation of tumor-related biomarkers(Ge et al., 2017; Pan et al., 2022; Wen et al., 2024). Regarding miRNA, Han et al. (2021) identified 22 differentially expressed miRNAs in patients with CCA using next-generation sequencing. The target genes of these miRNAs were enriched in the RAS-RAF-MEK-ERK, PI3K-AKT-mTOR, and p53 signaling pathways, suggesting their central role in the regulation of CCA cell proliferation, migration, and apoptosis. Specific miRNAs, such as miR-106b-5p and miR-23b-3p, were observed to be upregulated, and functional analysis showed that they promote tumor progression by enhancing tumor angiogenesis and anti-apoptotic abilities(Han et al., 2021). Furthermore, Pan et al. (2022) reported that miR-200c-3p in bile exhibited an AUC of 0.87 in CCA patients, which increased to 0.906 when combined with CA19-9. Elevated levels of miR-200c-3p were significantly associated with an increased risk of recurrence (hazard ratio [HR] = 2.78, $p< 0.01$), demonstrating its potential for clinical application(Pan, et al., 2022). Regarding lncRNA, Ge et al. (2017) identified significant upregulation of ENST00000588480.1 and ENST00000517758.1 in bile exosomes through RNA sequencing. These lncRNAs were associated with advanced TNM staging and could serve as independent prognostic biomarkers (HR = 3.14, $p = 0.002$) (Ge, et al., 2017). However, their specificity is relatively limited (58.9%), and further optimization of composite biomarkers is required to improve accuracy. Regarding circRNA, Wen et al. (2024) proposed a combined diagnostic model incorporating hsa-circ-0000367, hsa-circ-0021647, and hsa-circ-0000288 in bile and serum. The AUC for bile samples reached 0.947, significantly surpassing that of CA19-9 (AUC = 0.759)(Wen, et al., 2024).

7.2 Proteomics

In exosome research, the precise isolation of high-purity exosomes is paramount for data reliability. To address the challenges posed by high viscosity and protein contamination in bile samples, researchers incorporated chelating agents (e.g., EDEG) during ultracentrifugation. This innovative method effectively mitigated interference from contaminating proteins (e.g., albumin and immunoglobulins) and significantly improved the concentration and quality of exosomal RNA extraction (Fig. 3c)(Ikeda et al., 2021). The study identified 166 proteins that were significantly upregulated in CCA patients, with Claudin-3 (CLDN3) emerging as the most prominent diagnostic marker. Its concentration in a CCA cohort was significantly higher than that in a choledocholithiasis cohort (76.99 pg/mL vs. 25.51 pg/mL, $p = 0.0385$), with both sensitivity and specificity at 87.5%, and an AUC of 0.945. Immunohistochemical analyses also showed that Claudin-3 expression was significantly elevated in CCA tissue compared to normal biliary epithelium, further substantiating its potential as a diagnostic biomarker(Ikeda, et al., 2021). However, the study was limited by the small sample size of only 10 CCA patients and 10 choledocholithiasis patients and did not include other benign biliary diseases as a control group. This limits the generalizability of its findings and the reliability of its broader applicability.

Unedited

7.3 Metabolomics

Muraki et al. (2023) conducted a comprehensive lipidomic analysis of small extracellular vesicles (sEVs) in bile using LC-MS/MS and found that the level of phosphatidylcholine was significantly higher in patients with malignancy than in those with benign conditions (about 4.98-fold higher, $p = 0.037$). The diagnostic sensitivity was 71.4%, with a specificity of 100%, and an AUC of 0.857. Furthermore, the PC concentration in bile sEVs was measured using a commercial PC assay kit, and the results were consistent with those obtained through LC-MS/MS. This finding provides a simple tool for the clinical diagnosis of CCA, reducing reliance on sophisticated diagnostic instrumentation and laying the foundation for clinical implementation(Muraki et al., 2023). However, the study was limited by a small sample size (7 malignant and 8 benign cases), resulting in low statistical power, and further validation of the findings is required to confirm their reliability.

8 Conclusions and future perspectives

The early diagnosis of CCA is crucial for improving patient survival but the diagnostic efficacy of current serum biomarkers and imaging modalities remains suboptimal. Therefore, developing more sensitive and specific detection methods has become a research priority. Bile, as a bodily fluid derived directly from the biliary system, offers unique advantages for the early diagnosis of CCA owing to its accessibility and its abundance of tumor-related biomolecules. In recent years, the rapid development of omics technologies has provided significant support for research on bile samples, facilitating not only the elucidation of CCA molecular mechanisms but also the identification of a spectrum of high-value biomarkers (Table 1).

Despite multiple bile-based biomarkers across genomics, epigenomics, proteomics, metabolomics, and microbiomics reporting AUCs >0.90, only a subset have progressed beyond single-center case–control designs. Notable examples with prospective/external validation include next-generation sequencing of bile cfDNA for malignant biliary strictures, methylation panels in PSC-associated cohorts, and a combined bile–urine proteomic model that includes a prospective validation cohort. However, to our knowledge, no bile-based assay has yet completed multi-center prospective studies demonstrating clinical utility or achieved regulatory approval for routine use in CCA. Key obstacles include pre-analytical constraints inherent to bile sampling, such as reliance on ERCP or PTBD procedures, contrast-induced dilution, and inflammation-related matrix effects. In addition, interference from the disease spectrum, such as PSC, cholangitis, and choledocholithiasis, reduces real-world specificity. Methodological heterogeneity across platforms, cut-off values, normalization strategies, and quality control further limits reproducibility between centers. Another limitation is the lack of rigorous head-to-head comparisons against current diagnostic pathways, such as imaging and cytology/FISH, along with the absence of clinical-utility trials. In the future, concerted efforts should be directed towards expanding sample sizes, conducting multi-center prospective studies, optimizing detection technologies, and further elucidating the molecular mechanisms of CCA. Furthermore, the integration of multi-omics data will constitute a pivotal direction which, with the implementation of artificial intelligence and machine learning methodologies, is expected to enable precise subtyping of CCA, identification of therapeutic targets, and prognostic evaluation, thereby advancing the development of precision medicine and providing more efficient diagnostic and personalized therapeutic strategies for patients.

Uncorrected

Table 1 Biomarkers in human bile with CCA

Type	References	Sample size	Methods	Biomarkers	Results
Genomics	(Li et al., 2022b)	Bile and plasma samples from 13 CCA patients, with paired tumor tissue samples from 11 of them	NGS	/	The bile cfDNA concentration was significantly higher than in plasma (1918 vs. 63.1 ng/ml, p=0.0017); The gene mutation consistency between bile and tumor tissue reached 90%, significantly outperforming plasma (35%)
	(Lee et al., 2016)	24 bile samples (from patients with infiltrative CCA) and 17 tumor tissue samples (from patients with mass-forming CCA)	NGS	Infiltrative CCA: EGFR, FGFR1, ABL1; Mass-forming CCA: KRAS, TP53, APC	The DNA amount extracted from bile samples was 66.0 ± 84.7 μg, with high sequencing coverage depth; There were significant differences in the main mutated genes between the infiltrative and mass-forming types
	(Arechederra et al., 2022)	68 bile samples (26 benign, 9 uncertain, and 33 malignant); 54 blood samples (20 benign, 6 uncertain, and 28 malignant); 30 tumor tissue samples	NGS	KRAS, TP53, GNAS and EGFR	Sensitivity: 96.4% Specificity: 69.2% The sensitivity in uncertain cases reached 100%
	(Shen et al., 2022)	Bile samples from 4 patients with biliary tract cancer	3D-BCF	/	The cfDNA extraction efficiency of the 3D-BCF method was significantly improved, with an extraction amount of 300-1500 ng per milliliter of sample

Unedited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
Epigenomics	(Klump et al., 2003)	71 bile samples (23 from CCA, 5 from gallbladder cancer, 26 from common bile duct stones, 6 with no abnormalities, and 11 from primary sclerosing cholangitis)	Methylation-specific PCR (MSP)	p16INK4a p14ARF	The methylation rates of the two biomarkers in the malignant disease group were 52.2% (p16INK4a) and 46.2% (p14ARF); in the benign disease group, the methylation rates were much lower (6.3% for p16INK4a and 3.1% for p14ARF)
	(Zhang et al., 2010)	80 bile samples (23 from CCA, 21 from pancreatic cancer, and 36 from common bile duct stones)	MSP	Six-gene panel (DKK3, p16, SFRP2, DKK2, NPTX2, ppENK)	In the diagnosis of malignant bile duct stricture: Sensitivity: 77.27% Specificity: 77.78%
	(Shin et al., 2012)	77 bile samples from patients with extrahepatic cholangiocarcinoma (randomly divided into a training set of 40 cases, a validation set of 45 cases, and a test set of 40 cases)	MethyLight	Five-gene panel (CCND2, CDH13, GRIN2B, RUNX3, TWIST1)	In the test set: Sensitivity: 83% Specificity: 100%
	(Vedeld et al., 2021)	344 bile samples (205 from primary sclerosing cholangitis, 69 from non-PSC-related CCA, and 70 from benign cases)	ddPCR	Three-gene combined diagnosis (CDO1, SEPT9, VIM)	In patients with precancerous conditions and early-stage CCA: Sensitivity: 100% Specificity: 90% AUC = 0.87

Unedited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
Transcriptomics	(Gaetano et al., 2011)	18 bile samples (7 from CCA, 2 from gallbladder cancer, and 9 from common bile duct stones)	RT-qPCR and small RNA sequencing	miR-9	Sensitivity: 88.9% Specificity: 100% AUC = 0.91
	(Syn et al., 2015)	83 bile samples (19 from CCA, 52 from primary sclerosing cholangitis, and 12 from PSC/CCA); 83 serum samples (31 from CCA, 40 from primary sclerosing cholangitis, and 12 from normal controls)	High-throughput miRNA screening and qRT-PCR validation	Serum: miR-1281 miR-126 Bile: miR-1537 miR-412	Combined analysis of miR-1537 and CA19-9 in bile: Sensitivity: 73% Specificity: 93% AUC = 0.91
	(Han et al., 2020)	In the discovery group, there are 21 bile samples (11 from CCA and 10 from benign biliary diseases); In the validation group, there are 85 bile samples (37 from CCA and 48 from benign biliary diseases)	High-throughput miRNA screening and qRT-PCR validation	miR-30d-5p miR-92a-3p	miR-30d-5p： Sensitivity: 81.1% Specificity: 60.5% AUC = 0.730
Proteomics	(Ayaru et al., 2008)	102 bile samples (60 malignant and 42 benign); 30 tissue samples (15 malignant and 15 benign)	Automated immunofluorescence detection	Mcm5	Sensitivity: 66% Specificity: 97%

Unedited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
Proteomics	(Mulvenna et al., 2012)	29 bile samples (19 from CCA and 10 benign); 66 serum samples (30 from CCA, 13 benign, and 23 normal controls)	2D-PAGE and MALDI-TOF/TOF	SSP411	Bile: Sensitivity: 84% Specificity: 90% Serum: Sensitivity: 90% Specificity: 83.3%
	(Ren et al., 2019)	37 bile samples (29 from CCA and 8 normal); 136 tissue samples (127 from CCA and 9 normal)	iTRAQ and LC-MS/MS	TPD52 DNAJB1	Combined analysis of TPD52 and DNAJB1 in bile: Sensitivity: 93.1% Specificity: 88.7% AUC = 0.95
	(Matsuda et al., 2013)	58 bile samples (29 from CCA and 29 benign); 4 tumor tissue samples from CCA	Mass Spectrometry	WFA-positive L1CAM and MUC1	Combined diagnosis of WFA-positive L1CAM and MUC1: AUC = 0.93
	(Son et al., 2020)	23 bile samples (18 from CCA and 5 benign)	LC-MS	Igκ, APOE, ALB, SERPINC1, C4-A, C3 and A1AT	The abundance in bile from patients with extrahepatic cholangiocarcinoma increased by at least two-fold

Unedited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
Proteomics	(Farina et al., 2014)	29 bile samples (23 from pancreatic cancer, 4 from CCA , and 2 from ampullary cancer)	iTRAQ and LC-MS/MS	CEAM6	Combined analysis of bile CEAM6 and serum CA19-9: Sensitivity: 97% Specificity: 83% AUC = 0.96
	(Lukic et al., 2014)	13 bile samples (4 from CCA , 4 from pancreatic cancer, 3 from chronic pancreatitis, and 2 from gallstones)	iTRAQ and LC-MS/MS	OLF4 Syntenin-2 RAC1	The three main biomarkers (OLF4, Syntenin-2, and RAC1) showed significant differential expression in malignant bile duct stricture
	(Voigtländer et al., 2017)	Study group: 87 bile and 87 urine samples Validation group: 45 bile and 45 urine samples	CE-MS	Specific peptide biomarkers from the combined analysis of bile and urine proteomics	Study group: Sensitivity: 92% Specificity: 84% AUC = 0.96 Validation group: Sensitivity: 94% Specificity: 76%
	(Voigtländer et al., 2020)	128 bile and 128 urine samples (52 from CCA, 33 from primary sclerosing cholangitis, and 43 benign)	CE-MS	Bile biomarkers (ADAMTS4, trypsin, etc.) Urine biomarkers (albumin fragments, etc.)	Combined bile and urine biomarkers: Sensitivity: 92% Specificity: 84% AUC = 0.96

Unedited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
Metabolomics	(Song et al., 2018)	104 bile samples (30 from CCA, 17 from pancreatic cancer, and 57 benign)	LC-MS/MS	GCA TCDCA	The proportion of GCA is highest in CCA patients (35.6%), while the proportion of TCDCA was lowest (7.31%)
	(Zhang et al., 2023a)	49 bile samples (33 from pCCA and 16 from gallstones)	LC-MS/MS	Multiple bile metabolites, including glycocholic acid (C01921), tauroursodeoxycholic acid (C05466), etc	305 cationic modes and 246 anionic modes differential metabolites were selected
	(Navaneethan et al., 2014)	46 bile samples (8 from CCA, 17 from pancreatic cancer, 6 from primary sclerosing cholangitis, and 15 benign)	LC-ESI-MS/MS	ON-PC S-PC	Combined ON-PC and S-PC biomarkers: Sensitivity: 100% Specificity: 83.3% AUC = 0.91
	(Yang et al., 2025)	Exploratory cohort: 300 bile samples Validation cohort: 36 bile samples. Prospective cohort: 36 bile and plasma samples	NPELDI-MS	Six key metabolites: Glc, GCDCA, Lia, Pla, Sya, Aac	The BileMet model in the exploratory cohort: Sensitivity: 83.48% Specificity: 82.04% AUC = 0.891
	(Navaneethan et al., 2015)	32 bile samples (11 from primary sclerosing cholangitis complicated by CCA and 21 from PSC)	SIFT-MS	Acrylonitrile 3-Methylhexane Benzene	Sensitivity: 90.5% Specificity: 72.7% AUC = 0.89

Uncredited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
	(Gui et al., 2023)	200 bile samples (102 from pCCA and 98 benign)	GC-IMS	12 significantly different VOCs (4 upregulated and 8 downregulated)	SVM model: Sensitivity: 93.1% Specificity: 100% AUC = 0.966
Microbiomics	(Avilés-Jiménez et al., 2016)	200 bile samples (100 from EHC and 100 benign)	16S rRNA Gene Sequencing	Fusobacterium Prevotella Helicobacter pylori virulence genes (cagA, vacA)	Fusobacteria and Prevotella were significantly increased in EHC ($p<0.05$); the positivity rate of cagA and vacA in EHC was 50%, with the vacA positive rate significantly higher than in the BBP group ($p=0.0003$)
	(Chen et al., 2019)	68 bile samples (8 from dCCA, 44 from primary common bile duct stones, and 16 from recurrent common bile duct stones)	16S rRNA Gene Sequencing	Gemmatimonadetes Nitrospirae Staphylococcus Corynebacterium	The dCCA group was significantly higher than the bile duct stone group
	(Dangtakot et al., 2021)	60 bile samples (30 from CCA and 30 from common bile duct stones)	16S rRNA Gene Sequencing	Enterobacter Pseudomonas Cetobacterium Pyramidobacter	Enterobacter and Pseudomonas significantly increased in CCA, while Cetobacterium and Pyramidobacter were more abundant in CDL

Unedited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
Microbiomics	(Li et al., 2022a)	53 bile samples (14 from pCCA, 9 from dCCA, 8 from PC, and 22 from cholelithiasis)	16S rRNA Gene Sequencing	pCCA: Pseudomonas Sphingomonas Halomonas dCCA: Streptococcus Prevotella Halomonas PC: Pseudomonas Chloroplast Acinetobacter	The pCCA, dCCA, and PC groups each have unique microbial biomarkers. The cancer group exhibited a significant reduction in microbial diversity, whereas the CH group showed the highest diversity and richness
	(Zhang et al., 2023b)	42 bile samples (from CCA); 58 fecal samples (42 from CCA and 16 normal)	16S rRNA Gene Sequencing	The relative abundance ratio of Bifidobacterium to Klebsiella (Bifidobacterium/Klebsiella Ratio, B/K Ratio)	Gut microbiota: Sensitivity = 83.3% Specificity = 81.3% AUC = 0.87
	(Lee et al., 2024)	24 bile samples (11 from CCA and 13 benign)	16S rRNA Gene Sequencing and Nuclear Magnetic Resonance Analysis	Microbiome: Escherichia coli Metabolite: Isoleucine	Model combining microbial and metabolite features: Sensitivity: 91.0% Specificity: 87.0% AUC = 0.92

Unedited

Table 1. Cont.

Type	References	Sample size	Methods	Biomarkers	Results
Biliary Exosomes	(Han et al., 2021)	3 bile samples (2 from CCA and 1 benign)	High-throughput miRNA screening and qRT-PCR validation	miR-106b-5p miR-23b-3p	Upregulated in CCA patients
	(Pan et al., 2022)	100 bile samples (50 from CCA and 50 benign)	High-throughput miRNA screening and qRT-PCR validation	miR-200c-3p	Sensitivity: 83.3% Specificity: 86.7% AUC = 0.869
	(Ge et al., 2017)	91 bile samples (35 from CCA and 56 benign)	High-throughput RNA screening and qRT-PCR validation	ENST00000588480.1 ENST00000517758.1	Combination of two biomarkers: Sensitivity: 82.9% Specificity: 58.9% AUC = 0.709
	(Wen et al., 2024)	289 bile and 289 serum samples (164 from CCA and 125 benign for each)	High-throughput RNA screening and qRT-PCR validation	hsa-circ-0000367 hsa-circ-0021647 hsa-circ-0000288	Biliary exosome circRNA-based diagnostic model: AUC = 0.947
	(Ikeda et al., 2021)	20 bile samples (10 from CCA and 10 benign)	LC-MS	Claudin-3	Sensitivity: 87.5% Specificity: 87.5% AUC = 0.945
	(Muraki et al., 2023)	15 bile samples (7 from CCA and 8 benign)	LC-MS/MS	Phosphatidylcholine	Sensitivity: 71.4% Specificity: 100% AUC = 0.857

Unedited

Acknowledgments

This work is supported by Beijing Natural Science Foundation (No. L248022), National High Level Hospital Clinical Research Funding (No. 2022-PUMCH-B-128), CAMS Innovation Fund for Medical Sciences (CIFMS) (Nos. 2021-I2M-1-061 and 2021-I2M-1-003), CSCO-hengrui Cancer Research Fund (Nos. Y-HR2019-0239 and Y-HR2020MS-0415), Peking Union Medical College Hospital Outstanding Young Talent Development Program (No. UBJ10528).

Author contributions

Taifeng ZHU and Jiali XING wrote the original manuscript. Yongchang ZHENG and Haitao ZHAO conceived, designed, and discussed the work, and revised the manuscript. Ziyue HUANG, Shanshan WANG, Duo LI, Xinting SANG, Cong NING and Chengpei ZHU conducted comprehensive literature research and created all illustrations. All authors have read and approved the final manuscript.

Compliance with ethics guidelines

Taifeng ZHU, Jiali XING, Ziyue HUANG, Shanshan WANG, Duo LI, Xinting SANG, Cong NING, Chengpei ZHU, Yongchang ZHENG and Haitao ZHAO declare that they have no conflict of interest.

This review does not contain any studies with human or animal subjects performed by any of the authors.

References

Abe M, Kagara N, Miyake T, et al., 2019. Highly sensitive detection of sentinel lymph node metastasis of breast cancer by digital pcr for rassf1a methylation. *Oncology Reports*, https://doi.org/10.3892/or.2019.7363

Akhoundova D, Rubin MA, 2022. Clinical application of advanced multi-omics tumor profiling: Shaping precision oncology of the future. *Cancer Cell*, 40(9):920-938. https://doi.org/10.1016/j.ccell.2022.08.011

Altomare DF, Di Lena M, Porcelli F, et al., 2013. Exhaled volatile organic compounds identify patients with colorectal cancer. *Journal of British Surgery*, 100(1):144-150. https://doi.org/10.1002/bjs.8942

Apostolov E, Al-Soud WA, Nilsson I, et al., 2009. Helicobacter pyloriand otherhelicobacterspecies in gallbladder and liver of patients with chronic cholecystitis detected by immunological and molecular methods. *Scandinavian Journal of Gastroenterology*, 40(1):96-102. https://doi.org/10.1080/00365520410009546

Arechederra M, Rullán M, Amat I, et al., 2022. Next-generation sequencing of bile cell-free DNA for the early detection of patients with malignant biliary strictures. *Gut*, 71(6):1141-1151. https://doi.org/10.1136/gutjnl-2021-325178

Avilés-Jiménez F, Guitron A, Segura-López F, et al., 2016. Microbiota studies in the bile duct strongly suggest a role for helicobacter pylori in extrahepatic cholangiocarcinoma. *Clinical Microbiology and Infection*, 22(2):178.e111-178.e122. https://doi.org/10.1016/j.cmi.2015.10.008

Ayaru L, Stoeber K, Webster GJ, et al., 2008. Diagnosis of pancreaticobiliary malignancy by detection of minichromosome maintenance protein 5 in bile aspirates. *British Journal of Cancer*, 98(9):1548-1554. https://doi.org/10.1038/sj.bjc.6604342

Binda C, Gibiino G, Coluccio C, et al., 2022. Biliary diseases from the microbiome perspective: How microorganisms could change the approach to benign and malignant diseases. *Microorganisms*, 10(2) https://doi.org/10.3390/microorganisms10020312

Blumenthal RD, Leon E, Hansen HJ, et al., 2007. Expression patterns of ceacam5 and ceacam6 in primary and metastatic cancers. *BMC Cancer*, 7(1) https://doi.org/10.1186/1471-2407-7-2

Bombaci M, Pesce E, Torri A, et al., 2019. Novel biomarkers for primary biliary cholangitis to improve diagnosis and understand underlying regulatory mechanisms. *Liver International*, 39(11):2124-2135. https://doi.org/10.1111/liv.14128

Byrne MF, Gerke H, Mitchell RM, et al., 2004. Yield of endoscopic ultrasound-guided fine-needle aspiration of bile duct lesions. *Endoscopy*, 36(8):715-719. https://doi.org/10.1055/s-2004-825657

Cai Y, Wang Z, Guo S, et al., 2023. Detection, mechanisms, and therapeutic implications of oncometabolites. *Trends in Endocrinology & Metabolism*, 34(12):849-861. https://doi.org/10.1016/j.tem.2023.08.018

Cao P, Li F, Xiao Y, et al., 2022. Identification and validation of 7-lncrna signature of epigenetic disorders by comprehensive epigenetic analysis. *Disease Markers*, 2022:1-14. https://doi.org/10.1155/2022/5118444

Chang Y-C, Chen M-H, Yeh C-N, et al., 2020. Omics-based platforms: Current status and potential use for

Uncited

cholangiocarcinoma. *Biomolecules*, 10(10) https://doi.org/10.3390/biom10101377

Chen B, Fu SW, Lu L, et al., 2019. A preliminary study of biliary microbiota in patients with bile duct stones or distal cholangiocarcinoma. *BioMed Research International*, 2019:1-12. https://doi.org/10.1155/2019/1092563

Chen L, Yan H-X, Yang W, et al., 2009. The role of microrna expression pattern in human intrahepatic cholangiocarcinoma. *Journal of Hepatology*, 50(2):358-369. https://doi.org/10.1016/j.jhep.2008.09.015

Chin L, Pomerantz J, Depinho RA, 1998. The ink4a/arf tumor suppressor: One gene--two products--two pathways. *Trends Biochem Sci*, 23(8):291-296. https://doi.org/10.1016/s0968-0004(98)01236-5

Conrad K, Roggenbuck D, Laass MW, 2014. Diagnosis and classification of ulcerative colitis. *Autoimmunity Reviews*, 13(4-5):463-466. https://doi.org/10.1016/j.autrev.2014.01.028

Cote GA, Gore JA, Mcelyea SD, et al., 2014. A pilot study to develop a diagnostic test for pancreatic ductal adenocarcinoma based on differential expression of select mirna in plasma and bile. *American Journal of Gastroenterology*, 109(12):1942-1952. https://doi.org/10.1038/ajg.2014.331

Cui M, Cheng C, Zhang L, 2022. High-throughput proteomics: A methodological mini-review. *Laboratory Investigation*, 102(11):1170-1181. https://doi.org/10.1038/s41374-022-00830-7

Cummins JM, Velculescu VE, 2006. Implications of micro-rna profiling for cancer diagnosis. *Oncogene*, 25(46):6220-6227. https://doi.org/10.1038/sj.onc.1209914

Dangtakot R, Intuyod K, Ahooja A, et al., 2021. Profiling of bile microbiome identifies district microbial population between choledocholithiasis and cholangiocarcinoma patients. *Asian Pacific Journal of Cancer Prevention*, 22(1):233-240. https://doi.org/10.31557/apjcp.2021.22.1.233

De Meij TG, Larbi IB, Van Der Schee MP, et al., 2013. Electronic nose can discriminate colorectal carcinoma and advanced adenomas by fecal volatile biomarker analysis: Proof of principle study. *International Journal of Cancer*, 134(5):1132-1138. https://doi.org/10.1002/ijc.28446

De Vries AB, Van Der Heide F, Ter Steege RWF, et al., 2019. Limited diagnostic accuracy and clinical impact of single-operator peroral cholangioscopy for indeterminate biliary strictures. *Endoscopy*, 52(02):107-114. https://doi.org/10.1055/a-1061-7067

Duxbury MS, Ito H, Benoit E, et al., 2004. Ceacam6 is a determinant of pancreatic adenocarcinoma cellular invasiveness. *British Journal of Cancer*, 91(7):1384-1390. https://doi.org/10.1038/sj.bjc.6602113

Duxbury MS, Matros E, Clancy T, et al., 2005. Ceacam6 is a novel biomarker in pancreatic adenocarcinoma and panin lesions. *Annals of Surgery*, 241(3):491-496. https://doi.org/10.1097/01.sla.0000154455.86404.e9

Ebata T, Ercolani G, Alvaro D, et al., 2017. Current status on cholangiocarcinoma and gallbladder cancer. *Liver Cancer*, 6(1):59-65. https://doi.org/10.1159/000449493

Emmanuel C, Chiew Y-E, George J, et al., 2014. Genomic classification of serous ovarian cancer with adjacent borderline differentiates ras pathway and tp53-mutant tumors and identifies nras as an oncogenic driver. *Clinical Cancer Research*, 20(24):6618-6630. https://doi.org/10.1158/1078-0432.Ccr-14-1292

Ergin S, Kherad N, Alagoz M, 2022. Rna sequencing and its applications in cancer and rare diseases. *Molecular Biology Reports*, 49(3):2325-2333. https://doi.org/10.1007/s11033-021-06963-0

Fallone CA, Tran S, Semret M, et al., 2003. Helicobacter DNA in bile: Correlation with hepato‐biliary diseases. *Alimentary Pharmacology & Therapeutics*, 17(3):453-458. https://doi.org/10.1046/j.1365-2036.2003.01424.x

Farina A, Dumonceau JM, Frossard JL, et al., 2009. Proteomic analysis of human bile from malignant biliary stenosis induced by pancreatic cancer. *J Proteome Res*, 8(1):159-169. https://doi.org/10.1021/pr8004925

Farina A, Dumonceau J-M, Delhaye M, et al., 2011. A step further in the analysis of human bile proteome. *Journal of Proteome Research*, 10(4):2047-2063. https://doi.org/10.1021/pr200011b

Farina A, Dumonceau J-M, Antinori P, et al., 2014. Bile carcinoembryonic cell adhesion molecule 6 (ceam6) as a biomarker of malignant biliary stenoses. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(5):1018-1025. https://doi.org/10.1016/j.bbapap.2013.06.010

Ferrari Júnior AP, Lichtenstein DR, Slivka A, et al., 1994. Brush cytology during ercp for the diagnosis of biliary and pancreatic malignancies. *Gastrointest Endosc*, 40(2 Pt 1):140-145. https://doi.org/10.1016/s0016-5107(94)70155-5

Gaetano C, Shigehara K, Yokomuro S, et al., 2011. Real-time pcr-based analysis of the human bile micrornaome

Uncorrected

identifies mir-9 as a potential diagnostic biomarker for biliary tract cancer. *PLoS ONE*, 6(8) https://doi.org/10.1371/journal.pone.0023584

Gao L, Lin Y, Yue P, et al., 2023. Identification of a novel bile marker clusterin and a public online prediction platform based on deep learning for cholangiocarcinoma. *BMC Medicine*, 21(1) https://doi.org/10.1186/s12916-023-02990-9

Ge X, Wang Y, Nie J, et al., 2017. The diagnostic/prognostic potential and molecular functions of long non-coding rnas in the exosomes derived from the bile of human cholangiocarcinoma. *Oncotarget*, 8(41):69995-70005. https://doi.org/10.18632/oncotarget.19547

Giacomini I, Gianfanti F, Desbats MA, et al., 2021. Cholesterol metabolic reprogramming in cancer and its pharmacological modulation as therapeutic strategy. *Frontiers in Oncology*, 11 https://doi.org/10.3389/fonc.2021.682911

Glasbrenner B, Ardan M, Boeck W, et al., 1999. Prospective evaluation of brush cytology of biliary strictures during endoscopic retrograde cholangiopancreatography. *Endoscopy*, 31(9):712-717. https://doi.org/10.1055/s-1999-73

Gui X, Zhang X, Xin Y, et al., 2023. Identification and validation of volatile organic compounds in bile for differential diagnosis of perihilar cholangiocarcinoma. *Clinica Chimica Acta*, 541 https://doi.org/10.1016/j.cca.2023.117235

Gunasekaran G, Bekki Y, Lourdusamy V, et al., 2020. Surgical treatments of hepatobiliary cancers. *Hepatology*, 73(S1):128-136. https://doi.org/10.1002/hep.31325

Han HS, Kim MJ, Han J-H, et al., 2020. Bile-derived circulating extracellular mir-30d-5p and mir-92a-3p as potential biomarkers for cholangiocarcinoma. *Hepatobiliary & Pancreatic Diseases International*, 19(1):41-50. https://doi.org/10.1016/j.hbpd.2019.10.009

Han J-Y, Ahn KS, Kim YH, et al., 2021. Circulating micrornas as biomarkers in bile-derived exosomes of cholangiocarcinoma. *Annals of Surgical Treatment and Research*, 101(3) https://doi.org/10.4174/astr.2021.101.3.140

Han S, Guiberson ER, Li Y, et al., 2024. High-throughput identification of gut microbiome-dependent metabolites. *Nature Protocols*, 19(7):2180-2205. https://doi.org/10.1038/s41596-024-00980-6

Hanouneh IA, Zein NN, Cikach F, et al., 2014. The breathprints in patients with liver disease identify novel breath biomarkers in alcoholic hepatitis. *Clinical Gastroenterology and Hepatology*, 12(3):516-523. https://doi.org/10.1016/j.cgh.2013.08.048

Harding JJ, Khalil DN, Fabris L, et al., 2023. Rational development of combination therapies for biliary tract cancers. *Journal of Hepatology*, 78(1):217-228. https://doi.org/10.1016/j.jhep.2022.09.004

Herman JG, Merlo A, Mao L, et al., 1995. Inactivation of the cdkn2/p16/mts1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res*, 55(20):4525-4530.

Ieta K, Tanaka F, Utsunomiya T, et al., 2006. Ceacam6 gene expression in intrahepatic cholangiocarcinoma. *British Journal of Cancer*, 95(4):532-540. https://doi.org/10.1038/sj.bjc.6603276

Ikeda C, Haga H, Makino N, et al., 2021. Utility of claudin-3 in extracellular vesicles from human bile as biomarkers of cholangiocarcinoma. *Scientific Reports*, 11(1) https://doi.org/10.1038/s41598-021-81023-y

Jackson LE, Tomlinson JL, Alva-Ruiz R, et al., 2024. Metabolome-wide association identifies altered metabolites and metabolic pathways in the serum of patients with cholangiocarcinoma. *JHEP Reports*, 6(6) https://doi.org/10.1016/j.jhepr.2024.101068

Jahr S, Hentze H, Englisch S, et al., 2001. DNA fragments in the blood plasma of cancer patients: Quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res*, 61(4):1659-1665.

Joo I, Lee JM, Yoon JH, 2018. Imaging diagnosis of intrahepatic and perihilar cholangiocarcinoma: Recent advances and challenges. *Radiology*, 288(1):7-13. https://doi.org/10.1148/radiol.2018171187

Ke W, Zeng L, Hu Y, et al., 2018. Detection of early-stage extrahepatic cholangiocarcinoma in patients with biliary strictures by soluble b7-h4 in the bile. *Am J Cancer Res*, 8(4):699-707.

Kim T-M, 2013. Clinical applications of next-generation sequencing in colorectal cancers. *World Journal of Gastroenterology*, 19(40) https://doi.org/10.3748/wjg.v19.i40.6784

Kipp BR, Stadheim LM, Halling SA, et al., 2004. A comparison of routine cytology and fluorescence in situ

Unedited

hybridization for the detection of malignant bile duct strictures. *The American Journal of Gastroenterology*, 99(9):1675-1681. https://doi.org/10.1111/j.1572-0241.2004.30281.x

Klump B, Hsieh CJ, Dette S, et al., 2003. Promoter methylation of ink4a/arf as detected in bile-significance for the differential diagnosis in biliary disease. *Clin Cancer Res*, 9(5):1773-1778.

Kristiansen TZ, Bunkenborg J, Gronborg M, et al., 2004. A proteomic analysis of human bile. *Molecular & Cellular Proteomics*, 3(7):715-728. https://doi.org/10.1074/mcp.M400015-MCP200

Lankisch TO, Metzger J, Negm AA, et al., 2011. Bile proteomic profiles differentiate cholangiocarcinoma from primary sclerosing cholangitis and choledocholithiasis §δ. *Hepatology*, 53(3):875-884. https://doi.org/10.1002/hep.24103

Lapitz A, Azkargorta M, Milkiewicz P, et al., 2023. Liquid biopsy-based protein biomarkers for risk prediction, early diagnosis, and prognostication of cholangiocarcinoma. *Journal of Hepatology*, 79(1):93-108. https://doi.org/10.1016/j.jhep.2023.02.027

Lederer A-K, Rasel H, Kohnert E, et al., 2023. Gut microbiota in diagnosis, therapy and prognosis of cholangiocarcinoma and gallbladder carcinoma—a scoping review. *Microorganisms*, 11(9) https://doi.org/10.3390/microorganisms11092363

Lee CH, Wang HE, Seo SY, et al., 2016. Cancer related gene alterations can be detected with next-generation sequencing analysis of bile in diffusely infiltrating type cholangiocarcinoma. *Experimental and Molecular Pathology*, 101(1):150-156. https://doi.org/10.1016/j.yexmp.2016.07.010

Lee J, Kim H, Park J-S, 2024. Beyond the bile: Exploring the microbiome and metabolites in cholangiocarcinoma. *Life*, 14(6) https://doi.org/10.3390/life14060698

Li J, Chen Q, Guo L, et al., 2022a. In situ detecting lipids as potential biomarkers for the diagnosis and prognosis of intrahepatic cholangiocarcinoma. *Cancer Management and Research*, Volume 14:2903-2912. https://doi.org/10.2147/cmar.S357000

Li Z, Chu J, Su F, et al., 2022b. Characteristics of bile microbiota in cholelithiasis, perihilar cholangiocarcinoma, distal cholangiocarcinoma, and pancreatic cancer. *Am J Transl Res*, 14(5):2962-2971.

Li Z, Liu Y, Fu J, et al., 2022c. Bile is a reliable and valuable source to study cfdna in biliary tract cancers. *Frontiers in Oncology*, 12 https://doi.org/10.3389/fonc.2022.961939

Liang Z, Liu X, Zhang Q, et al., 2016. Diagnostic value of micrornas as biomarkers for cholangiocarcinoma. *Digestive and Liver Disease*, 48(10):1227-1232. https://doi.org/10.1016/j.dld.2016.07.006

Liu Z, Pang B, Wang Y, et al., 2024. Advances of new extracellular vesicle isolation and detection technologies in cancer diagnosis. *Small*, 21(4) https://doi.org/10.1002/smll.202405872

Lukic N, Visentin R, Delhaye M, et al., 2014. An integrated approach for comparative proteomic analysis of human bile reveals overexpressed cancer-associated proteins in malignant biliary stenosis. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(5):1026-1033. https://doi.org/10.1016/j.bbapap.2013.06.023

Luo H, Zhao Q, Wei W, et al., 2020. Circulating tumor DNA methylation profiles enable early diagnosis, prognosis prediction, and screening for colorectal cancer. *Sci Transl Med*, 12(524) https://doi.org/10.1126/scitranslmed.aax7533

Matsuda A, Kuno A, Matsuzaki H, et al., 2013. Glycoproteomics-based cancer marker discovery adopting dual enrichment with wisteria floribunda agglutinin for high specific glyco-diagnosis of cholangiocarcinoma. *Journal of Proteomics*, 85:1-11. https://doi.org/10.1016/j.jprot.2013.04.017

Maury E, Hashizume R, 2017. Epigenetic modification in chromatin machinery and its deregulation in pediatric brain tumors: Insight into epigenetic therapies. *Epigenetics*, 12(5):353-369. https://doi.org/10.1080/15592294.2016.1278095

Meng F, Henson R, Lang M, et al., 2006. Involvement of human micro-rna in growth and response to chemotherapy in human cholangiocarcinoma cell lines. *Gastroenterology*, 130(7):2113-2129. https://doi.org/10.1053/j.gastro.2006.02.057

Merlo A, Herman JG, Mao L, et al., 1995. 5' cpg island methylation is associated with transcriptional silencing of the tumour suppressor p16/cdkn2/mts1 in human cancers. *Nat Med*, 1(7):686-692. https://doi.org/10.1038/nm0795-686

Uncorrected

Metzger J, Negm AA, Plentz RR, et al., 2013. Urine proteomic analysis differentiates cholangiocarcinoma from primary sclerosing cholangitis and other benign biliary disorders. *Gut*, 62(1):122-130. https://doi.org/10.1136/gutjnl-2012-302047

Miceli RT, Chen TY, Nose Y, et al., 2024. Extracellular vesicles, rna sequencing, and bioinformatic analyses: Challenges, solutions, and recommendations. *Journal of Extracellular Vesicles*, 13(12) https://doi.org/10.1002/jev2.70005

Mima K, Nakagawa S, Sawayama H, et al., 2017. The microbiome and hepatobiliary-pancreatic cancers. *Cancer Letters*, 402:9-15. https://doi.org/10.1016/j.canlet.2017.05.001

Miyabe K, Chandrasekhara V, Wongjarupong N, et al., 2022. Potential role of inflammation-promoting biliary microbiome in primary sclerosing cholangitis and cholangiocarcinoma. *Cancers*, 14(9) https://doi.org/10.3390/cancers14092120

Mulvenna J, Shen J, Wang W, et al., 2012. Comparative proteomic profiling of human bile reveals ssp411 as a novel biomarker of cholangiocarcinoma. *PLoS ONE*, 7(10) https://doi.org/10.1371/journal.pone.0047476

Muraki R, Morita Y, Ida S, et al., 2023. Phosphatidylcholine in bile‐derived small extracellular vesicles as a novel biomarker of cholangiocarcinoma. *Cancer Medicine*, 12(12):13007-13018. https://doi.org/10.1002/cam4.5973

Nakaoka T, Saito Y, Saito H, 2017. Aberrant DNA methylation as a biomarker and a therapeutic target of cholangiocarcinoma. *International Journal of Molecular Sciences*, 18(6) https://doi.org/10.3390/ijms18061111

Navaneethan U, Gutierrez NG, Venkatesh PGK, et al., 2014a. Lipidomic profiling of bile in distinguishing benign from malignant biliary strictures: A single-blinded pilot study. *American Journal of Gastroenterology*, 109(6):895-902. https://doi.org/10.1038/ajg.2014.60

Navaneethan U, Parsi MA, Gutierrez NG, et al., 2014b. Volatile organic compounds in bile can diagnose malignant biliary strictures in the setting of pancreatic cancer: A preliminary observation. *Gastrointestinal Endoscopy*, 80(6):1038-1045. https://doi.org/10.1016/j.gie.2014.04.016

Navaneethan U, Parsi MA, Lourdusamy V, et al., 2015. Volatile organic compounds in bile for early diagnosis of cholangiocarcinoma in patients with primary sclerosing cholangitis: A pilot study. *Gastrointestinal Endoscopy*, 81(4):943-949.e941. https://doi.org/10.1016/j.gie.2014.09.041

Nehls O, Gregor M, Klump B, 2004. Serum and bile markers for cholangiocarcinoma. *Semin Liver Dis*, 24(2):139-154. https://doi.org/10.1055/s-2004-828891

Nehring P, Przybyłkowski A, Podgórski PB, 2023. Bile composition - which components can have clinical significance? A review of the literature. *Journal of Gastrointestinal and Liver Diseases*, 32(1):110-117. https://doi.org/10.15403/jgld-4634

Padthaisong S, Phetcharaburanin J, Klanrit P, et al., 2021. Integration of global metabolomics and lipidomics approaches reveals the molecular mechanisms and the potential biomarkers for postoperative recurrence in early-stage cholangiocarcinoma. *Cancer & Metabolism*, 9(1) https://doi.org/10.1186/s40170-021-00266-5

Pan Y, Shao S, Sun H, et al., 2022. Bile-derived exosome noncoding rnas as potential diagnostic and prognostic biomarkers for cholangiocarcinoma. *Frontiers in Oncology*, 12 https://doi.org/10.3389/fonc.2022.985089

Panea RI, Love CL, Shingleton JR, et al., 2019. The whole-genome landscape of burkitt lymphoma subtypes. *Blood*, 134(19):1598-1607. https://doi.org/10.1182/blood.2019001880

Ptashkin RN, Mandelker DL, Coombs CC, et al., 2018. Prevalence of clonal hematopoiesis mutations in tumor-only clinical genomic profiling of solid tumors. *JAMA Oncology*, 4(11) https://doi.org/10.1001/jamaoncol.2018.2297

Quelle DE, Zindy F, Ashmun RA, et al., 1995. Alternative reading frames of the ink4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. *Cell*, 83(6):993-1000. https://doi.org/10.1016/0092-8674(95)90214-7

Razumilava N, Gores GJ, 2013. Classification, diagnosis, and management of cholangiocarcinoma. *Clinical Gastroenterology and Hepatology*, 11(1):13-21.e11. https://doi.org/10.1016/j.cgh.2012.09.009

Ren H, Luo M, Chen J, et al., 2019. Identification of tpd52 and dnajb1 as two novel bile biomarkers for cholangiocarcinoma by itraq-based quantitative proteomics analysis. *Oncology Reports*,

Uncited

https://doi.org/10.3892/or.2019.7387

Rodrigues PM, Olaizola P, Paiva NA, et al., 2021. Pathogenesis of cholangiocarcinoma. *Annual Review of Pathology: Mechanisms of Disease*, 16(1):433-463. https://doi.org/10.1146/annurev-pathol-030220-020455

Rösch T, Meining A, Frühmorgen S, et al., 2002. A prospective comparison of the diagnostic accuracy of ercp, mrcp, ct, and eus in biliary strictures. *Gastrointestinal Endoscopy*, 55(7):870-876. https://doi.org/10.1067/mge.2002.124206

Russell DW, 2003. The enzymes, regulation, and genetics of bile acid synthesis. *Annual Review of Biochemistry*, 72(1):137-174. https://doi.org/10.1146/annurev.biochem.72.121801.161712

Sager M, Yeat NC, Pajaro-Van Der Stadt S, et al., 2015. Transcriptomics in cancer diagnostics: Developments in technology, clinical research and commercialization. *Expert Review of Molecular Diagnostics*, 15(12):1589-1603. https://doi.org/10.1586/14737159.2015.1105133

Satam H, Joshi K, Mangrolia U, et al., 2023. Next-generation sequencing technology: Current trends and advancements. *Biology*, 12(7) https://doi.org/10.3390/biology12070997

Segura‐López FK, Avilés‐Jiménez F, Güitrón‐Cantú A, et al., 2015. Infection with helicobacter bilis but not helicobacter hepaticus was associated with extrahepatic cholangiocarcinoma. *Helicobacter*, 20(3):223-230. https://doi.org/10.1111/hel.12195

Shen N, Zhu B, Zhang W, et al., 2022. Comprehensive evaluation and application of a novel method to isolate cell-free DNA derived from bile of biliary tract cancer patients. *Frontiers in Oncology*, 12 https://doi.org/10.3389/fonc.2022.891917

Shin S-H, Lee K, Kim B-H, et al., 2012. Bile-based detection of extrahepatic cholangiocarcinoma with quantitative DNA methylation markers and its high sensitivity. *The Journal of Molecular Diagnostics*, 14(3):256-263. https://doi.org/10.1016/j.jmoldx.2012.01.014

Shu L, Li X, Liu Z, et al., 2024. Bile exosomal mir-182/183-5p increases cholangiocarcinoma stemness and progression by targeting hpgd and increasing pge2 generation. *Hepatology*, 79(2):307-322. https://doi.org/10.1097/hep.0000000000000437

Sinakos E, Saenger AK, Keach J, et al., 2011. Many patients with primary sclerosing cholangitis and increased serum levels of carbohydrate antigen 19-9 do not have cholangiocarcinoma. *Clinical Gastroenterology and Hepatology*, 9(5):434-439.e431. https://doi.org/10.1016/j.cgh.2011.02.007

Siravegna G, Marsoni S, Siena S, et al., 2017. Integrating liquid biopsies into the management of cancer. *Nature Reviews Clinical Oncology*, 14(9):531-548. https://doi.org/10.1038/nrclinonc.2017.14

Son KH, Ahn CB, Kim HJ, et al., 2020. Quantitative proteomic analysis of bile in extrahepatic cholangiocarcinoma patients. *Journal of Cancer*, 11(14):4073-4080. https://doi.org/10.7150/jca.40964

Song J, Li Y, Bowlus CL, et al., 2019. Cholangiocarcinoma in patients with primary sclerosing cholangitis (psc): A comprehensive review. *Clinical Reviews in Allergy & Immunology*, 58(1):134-149. https://doi.org/10.1007/s12016-019-08764-7

Song W-S, Park H-M, Ha JM, et al., 2018. Discovery of glycocholic acid and taurochenodeoxycholic acid as phenotypic biomarkers in cholangiocarcinoma. *Scientific Reports*, 8(1) https://doi.org/10.1038/s41598-018-29445-z

Syn W-K, Voigtländer T, Gupta SK, et al., 2015. Micrornas in serum and bile of patients with primary sclerosing cholangitis and/or cholangiocarcinoma. *Plos One*, 10(10) https://doi.org/10.1371/journal.pone.0139305

Tang H, Zhang Y, Wu Y, et al., 2024. The emerging era of molecular medicine. *ACS Nano*, 18(45):30911-30918. https://doi.org/10.1021/acsnano.4c07969

Thomas C, Pellicciari R, Pruzanski M, et al., 2008. Targeting bile-acid signalling for metabolic diseases. *Nature Reviews Drug Discovery*, 7(8):678-693. https://doi.org/10.1038/nrd2619

Tomacha J, Dokduang H, Padthaisong S, et al., 2021. Targeting fatty acid synthase modulates metabolic pathways and inhibits cholangiocarcinoma cell progression. *Frontiers in Pharmacology*, 12 https://doi.org/10.3389/fphar.2021.696961

Tyson GL, El-Serag HB, 2011. Risk factors for cholangiocarcinoma. *Hepatology*, 54(1):173-184. https://doi.org/10.1002/hep.24351

Valle JW, Kelley RK, Nervi B, et al., 2021. Biliary tract cancer. *The Lancet*, 397(10272):428-444.

Unedited

https://doi.org/10.1016/s0140-6736(21)00153-7

Vassiliou E, Farias-Pereira R, 2023. Impact of lipid metabolism on macrophage polarization: Implications for inflammation and tumor immunity. *International Journal of Molecular Sciences*, 24(15) https://doi.org/10.3390/ijms241512032

Vedeld HM, Grimsrud MM, Andresen K, et al., 2021. Early and accurate detection of cholangiocarcinoma in patients with primary sclerosing cholangitis by methylation markers in bile. *Hepatology*, 75(1):59-73. https://doi.org/10.1002/hep.32125

Voigtländer T, Metzger J, Schönemeier B, et al., 2017. A combined bile and urine proteomic test for cholangiocarcinoma diagnosis in patients with biliary strictures of unknown origin. *United European Gastroenterology Journal*, 5(5):668-676. https://doi.org/10.1177/2050640616687836

Voigtländer T, Metzger J, Husi H, et al., 2020. Bile and urine peptide marker profiles: Access keys to molecular pathways and biological processes in cholangiocarcinoma. *Journal of Biomedical Science*, 27(1) https://doi.org/10.1186/s12929-019-0599-5

Wahlström A, Sayin Sama i, Marschall H-U, et al., 2016. Intestinal crosstalk between bile acids and microbiota and its impact on host metabolism. *Cell Metabolism*, 24(1):41-50. https://doi.org/10.1016/j.cmet.2016.05.005

Wang D, Shen Y, Qian H, et al., 2024. Emerging advanced approaches for liquid biopsy: In situ nucleic acid assays of extracellular vesicles. *Theranostics*, 14(19):7309-7332. https://doi.org/10.7150/thno.102437

Wang Z, Zhu H, Xiong W, 2023. Advances in mass spectrometry-based multi-scale metabolomic methodologies and their applications in biological and clinical investigations. *Science Bulletin*, 68(19):2268-2284. https://doi.org/10.1016/j.scib.2023.08.047

Wannhoff A, Rupp C, Friedrich K, et al., 2015. Inflammation but not biliary obstruction is associated with carbohydrate antigen 19-9 levels in patients with primary sclerosing cholangitis. *Clinical Gastroenterology and Hepatology*, 13(13):2372-2379. https://doi.org/10.1016/j.cgh.2015.07.014

Wasenang W, Chaiyarit P, Proungvitaya S, et al., 2019. Serum cell-free DNA methylation of opcml and hoxd9 as a biomarker that may aid in differential diagnosis between cholangiocarcinoma and other biliary diseases. *Clinical Epigenetics*, 11(1) https://doi.org/10.1186/s13148-019-0634-0

Wen N, Peng D, Xiong X, et al., 2024. Cholangiocarcinoma combined with biliary obstruction: An exosomal circrna signature for diagnosis and early recurrence monitoring. *Signal Transduction and Targeted Therapy*, 9(1) https://doi.org/10.1038/s41392-024-01814-3

Yang S, Fu J, Qin W, et al., 2025. Bile metabolic fingerprints distinguish biliary tract cancer from benign biliary diseases. *Hepatology*, 81(2):476-490. https://doi.org/10.1097/hep.0000000000000957

Ye C, Dong C, Lin Y, et al., 2023. Interplay between the human microbiome and biliary tract cancer: Implications for pathogenesis and therapy. *Microorganisms*, 11(10) https://doi.org/10.3390/microorganisms11102598

Zhang D, Zhao G, Sun W, et al., 2023a. Bile metabolites as diagnostic biomarkers for perihilar cholangiocarcinoma. *Scientific Reports*, 13(1) https://doi.org/10.1038/s41598-023-27603-6

Zhang N, Zhu W, Zhang S, et al., 2023b. A novel bifidobacterium/klebsiella ratio in characterization analysis of the gut and bile microbiota of cca patients. *Microbial Ecology*, 87(1) https://doi.org/10.1007/s00248-023-02318-3

Zhang Y, Yang B, Du Z, et al., 2010. Identification and validation of specific methylation profile in bile for differential diagnosis of malignant biliary stricture. *Clinical Biochemistry*, 43(16-17):1340-1344. https://doi.org/10.1016/j.clinbiochem.2010.08.013

Zhang Y, Fonslow BR, Shan B, et al., 2013. Protein analysis by shotgun/bottom-up proteomics. *Chemical Reviews*, 113(4):2343-2394. https://doi.org/10.1021/cr3003533

Under review